

DOCUMENT RESUME

ED 077 305

FL 004 424

AUTHOR Clausing, Gerhard; And Others
TITLE Individualized German Instruction at the College Level--A Second Evaluation.
PUB DATE 73
NOTE 33p.
EDRS PRICE MF-\$0.65 HC-\$3.29
DESCRIPTORS *Achievement Tests; Aptitude Tests; Curriculum Development; Diagrams; Enrollment Influences; Enrollment Rate; *German; *Individualized Instruction; Instructional Innovation; *Language Instruction; Language Learning Levels; Pretesting; Reading Skills; Speech Skills; *Statistical Data; Student Evaluation; Tables (Data); Test Results; Universities; Writing Skills
IDENTIFIERS Berkeley; University of California

ABSTRACT

This follow-up report to 'Individualized German Instruction at the College Level--A First Appraisal,' by Gerhard Clausing et al in "Foreign Language Annals," October 1972, employs a more rigorous analysis of pretest and posttest data than the first study and focuses specifically on test results involving the individualized and basic courses during the second quarter of the second year of the program. The evaluation of student achievement shows significantly better performance by individualized instruction students in reading and writing and equal oral performance compared with the basic course students. The findings confirm the conclusions of the first study and substantiate them with more rigorous statistical evidence. The data gathered over two years from a substantial student population are proposed as a demonstration of the basic soundness of individualized foreign language instruction.
(Author/RS)

ED 077305

INDIVIDUALIZED GERMAN INSTRUCTION AT THE COLLEGE LEVEL--A SECOND EVALUATION

Gerhard Clausing
Klaus A. Mueller
Anneliese Stein

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRO-
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN-
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT
OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY.

I. Introduction

A description and first evaluation of the Individualized Instruction Program in German (hereafter referred to as I.I.) at Berkeley was presented after the program had been in operation for one year. That study¹ contains the description of the teaching materials and methods of instruction for I.I., and contrasts this new course with the other courses offered by the department.² These courses, together with I.I., continue to be offered. In order to continue to measure the relative success of individualized instruction, this second evaluation was undertaken.

The necessity for replication of research is evident. This seems especially important in the case of studies involving instructional innovation. The second study utilizes pretests used in much previous research. In addition, a greater number and variety of posttests were used and administered in a more systematic manner. The second study also employs a more differentiated and rigorous analysis of pretest and posttest data than the first study. As a consequence, the second study is in part a replication of, and in part different from, the first evaluation. It should also be noted that the data reported in both studies was gathered from a population of students who determined the course in which they would enroll,

¹Gerhard Clausing, Klaus A. Mueller and Wilfried M. Voge, "Individualized German Instruction at the College Level--A First Appraisal," Foreign Language Annals, Vol. 6, No. 1, October 1972.

²The Basic Course, the Intensive Course, and the Course Emphasizing Reading.

FL 004 424

and that only thereafter sampling procedures were randomized.

Since the course materials and instructional modes were described in detail in the first study, we will not repeat this information here. We shall focus our report on the descriptions and results of the pretests and posttests which were employed during the second quarter of the second year of the program. Specifically we shall deal with tests and test results involving the I.I. and Basic Course (hereafter referred to as B.C.) students, since differences in course procedures, materials and performance are perhaps of greatest interest in these two groups.

In addition to the use of more suitable pretests and better controlled posttests we were able to match students in both courses more equally numerically than was possible in the first study. This resulted in our obtaining a substantially greater degree of confidence in the analysis of student achievement.

II. Test Measures and Data Analysis

A. Pretests

The students of the B.C. and I.I. were given the MLAT³ and the Seashore⁴ tests during enrollment and during the first day of classes.

1. MLAT

A t-test was computed for the MLAT pretest scores. The I.I. group sample was 137, the B.C. sample was 93. The t-value obtained was .7736, which is clearly not significant. Therefore we considered the two groups equivalent on the pretest scores of this measure for the purposes of our study.

³John B. Carroll and Stanley M. Sapon, Modern Language Aptitude Test, Form A (New York: The Psychological Corporation, 1958).

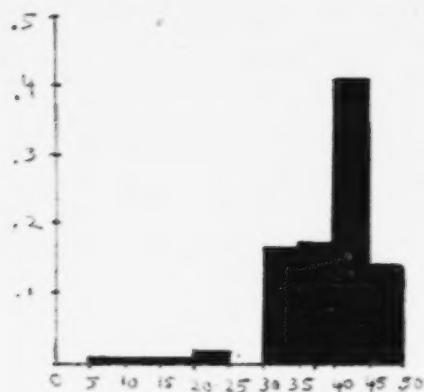
⁴Carl E. Seashore, Don Lewis and Joseph G. Saetveit, Seashore Measures of Musical Talents, Series A (New York: The Psychological Corporation, 1939).

2. Seashore

As may be observed in the histograms of the Seashore test results presented below (Figures 1-6), the B.C. and I.I. students were almost identical and there were no significant deviations between the two groups. Because of this lack of significant differences of both groups we considered the two groups equivalent on this pre-test for the purposes of our study.

Figure 1: Seashore--Pitch

B.C.



I.I.

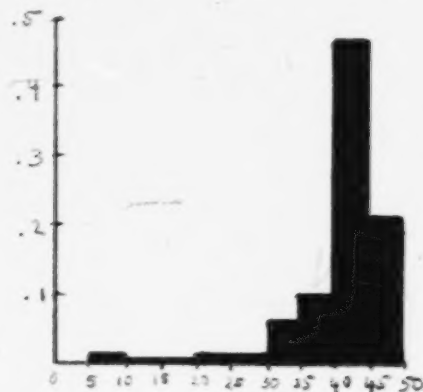
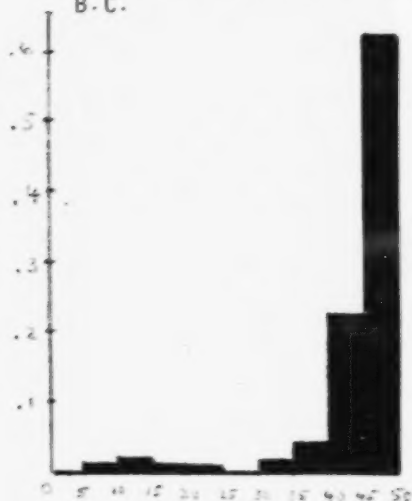


Figure 2: Seashore--Loudness

B.C.



I.I.

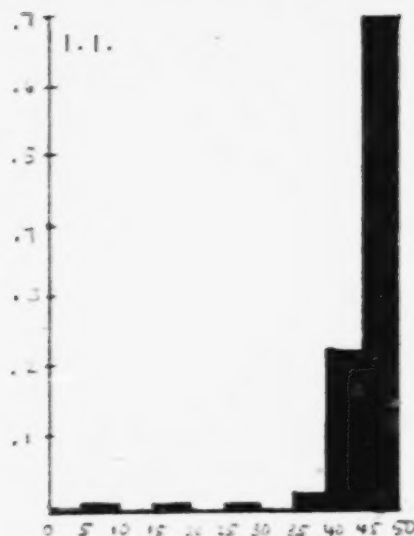


Figure 3: Seashore--Rhythm
B.C.

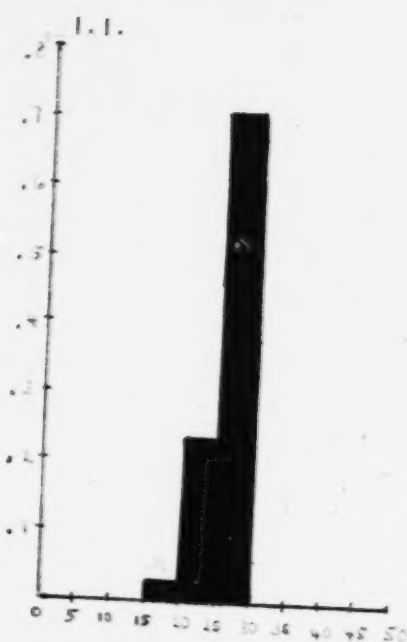
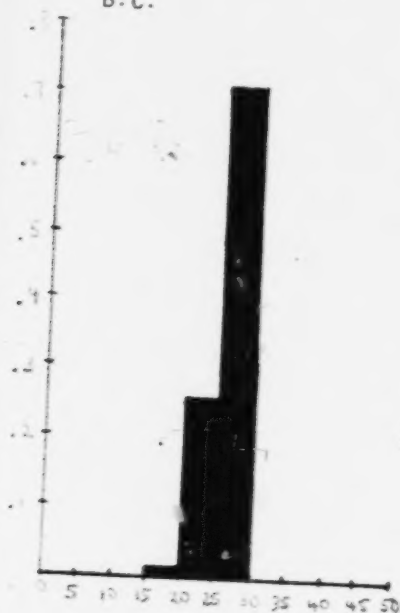


Figure 4: Seashore--Time
B.C.

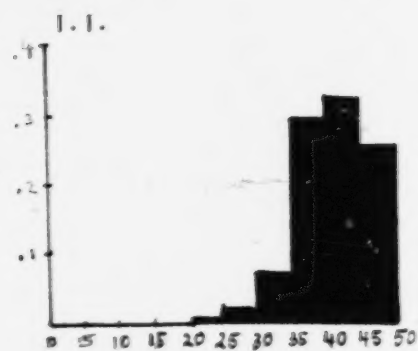
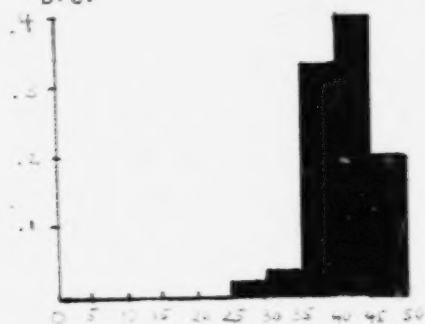


Figure 5: Seashore--Timbre
B.C.

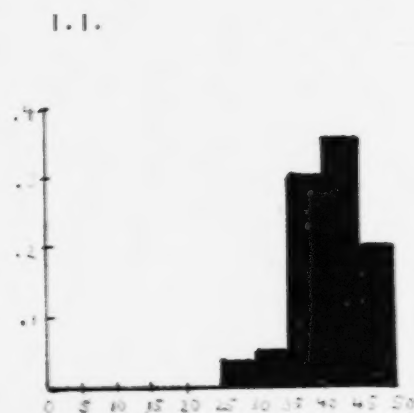
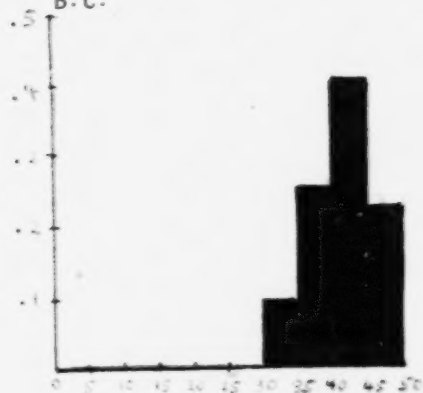
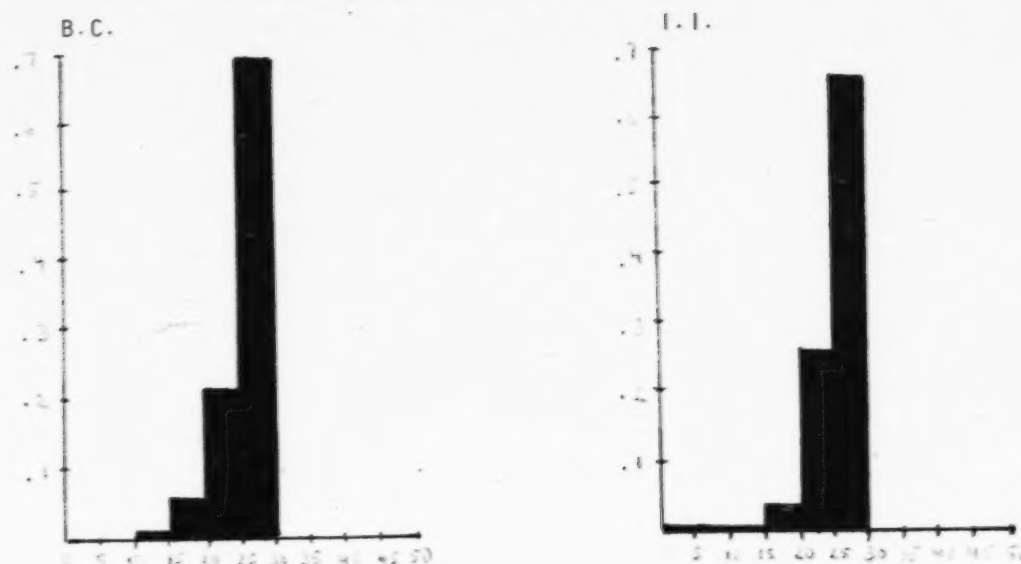


Figure 6: Seashore--Tonal Memory



B. Posttests

For each comparison the pretest score MLAT and the posttest scores L_1 , L_2 , $\bar{U}_I$ and $\bar{U}_{II}$ (consisting of total number of minus points) were compared. $\bar{U}_I$ denotes the average of U_2 , U_3 and U_4 , which are thirty-minute I.I. tests taken after each respective unit in the first quarter of German study. L_1 is the three-hour final B.C. test given regularly at the end of the first quarter of German. $\bar{U}_{II}$ is the average of U_6 , U_7 and U_8 , thirty-minute I.I. tests taken after each respective unit in the second quarter of German study. L_2 is the three-hour final B.C. test given regularly at the end of the second quarter of German study. All of these tests were administered to randomly selected B.C. and I.I. sample groups at appropriate times.

It was deemed necessary to see whether the MLAT scores would have an effect on the posttest scores. For this purpose the regression lines were computed and

fitted into the graphs. For each student a dot was plotted, representing his pretest score and his posttest score; then the regression line was fitted through the points estimating the ideal line which expresses the relation between the pretest score and the posttest score. This was done separately for the B.C. and I.I.

group in Level 1 (first quarter of German) and Level 2 (second quarter of German).

Then the F-test was applied. The F-test compares the spread in one population (I.I. students) against the spread in another population (B.C. students).

A t-test was then done for each comparison. The t-test compares the difference in means in populations relative to the spread within the populations.

Comparison 1:

The two groups, I.I. and B.C., were compared on the basis of their test scores in L_1 . There were 41 observations from B.C. and 13 observations from I.I.

a. The ordinary t-test (without regard to pretest scores) yields a value of +2.66 which is significant at level .005 (one-sided test).

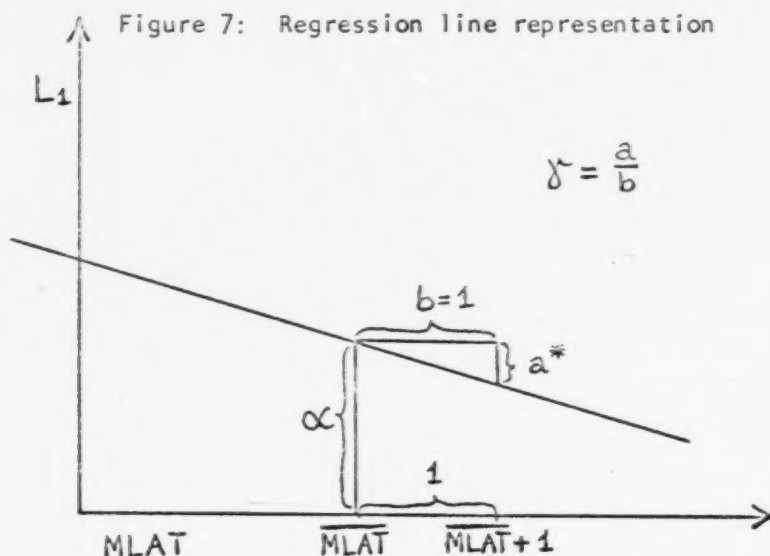
b. Taking into consideration the pretest score MLAT, we tried to fit a regression line to the observed values of $(MLAT, L_1)$ in the I.I. group, and the same in group B.C. This was done to explain some of the variation in the L_1 scores by the variations in the MLAT-scores of the involved students.

For each student a dot was plotted representing his MLAT and L_1 score. The regression line was then computed and plotted. This was done separately for the B.C. and I.I. group.

If the equation of the regression line is

$$L_1 = \alpha + \beta (MLAT - \overline{MLAT})$$

then γ represents the slope of the line, α is the intercept of the line at $MLAT = \overline{MLAT}$ ($\overline{MLAT}$ is the average of $MLAT$ -scores for the respective group under consideration).



*here a is negative

The following values for α and γ were computed as was s^2 (the residual sum of squares divided by the degrees of freedom), which measures the variation of the data around the fitted line.

$$\begin{aligned} \hat{\alpha} &= 23.38 \\ \hat{\gamma} &= -.322 \\ s^2 &= 50.899 \end{aligned}$$

$$\begin{aligned} \hat{\alpha} &= 36.44 \\ \hat{\gamma} &= -.396 \\ s^2 &= 236.718 \end{aligned}$$

$\hat{}$ denotes that a specific value is attached to the symbol.

c. The F-test for comparing the two variations around their respective regression lines shows a significant difference (at level .005), the F-value being 4.56 (degrees of freedom: 39 and 11 respectively).

d. The t-test for equality of slopes gives a t-value of .35 (50 degrees of freedom), which is not significant.

e. Considering the slopes equal for I.I. and B.C., one can test for differences in the $\bar{X}$ values and here the t-value is +2.77 (51 degrees of freedom), which is clearly significant at level .005.

Figure 8

LEVEL 1: Long Tests (L₁)- B.C.

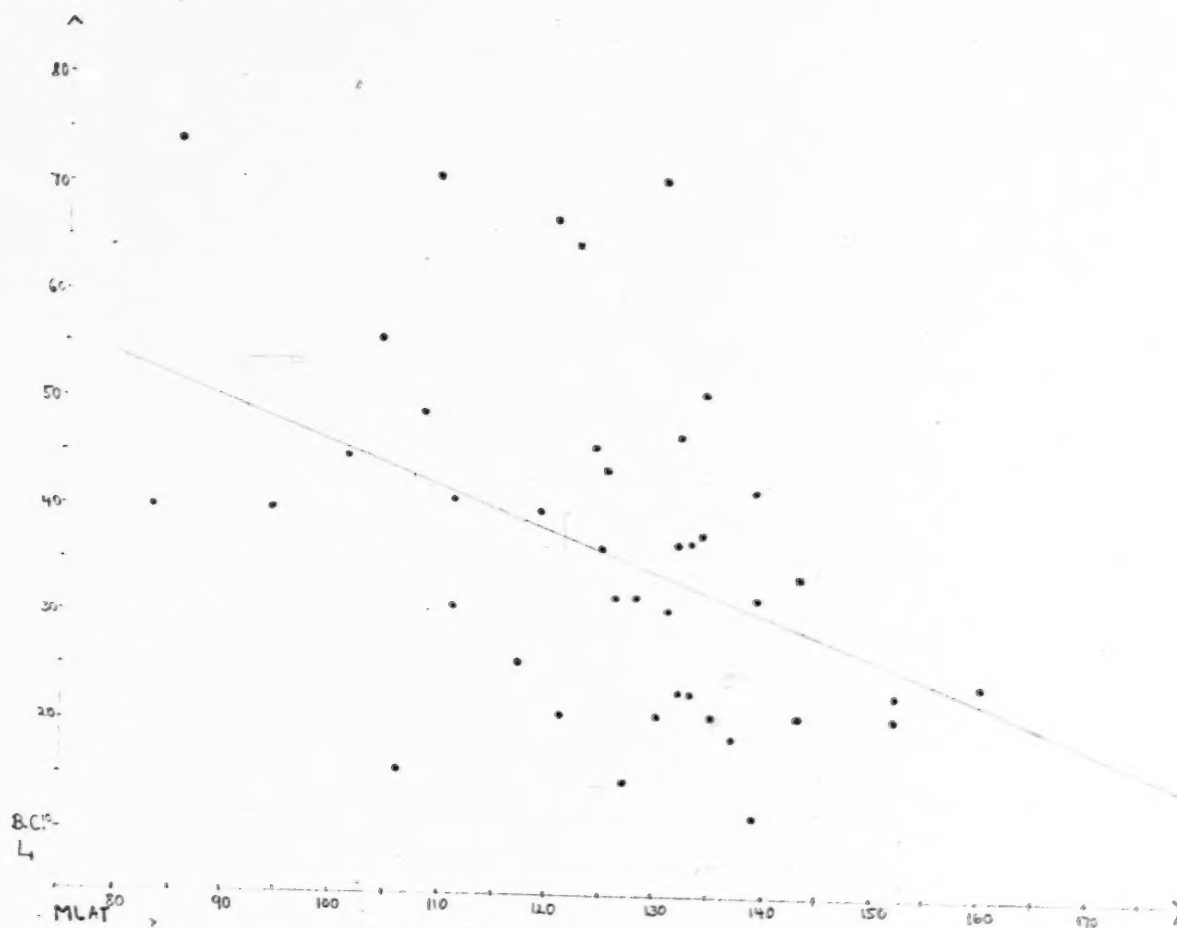


Figure 9

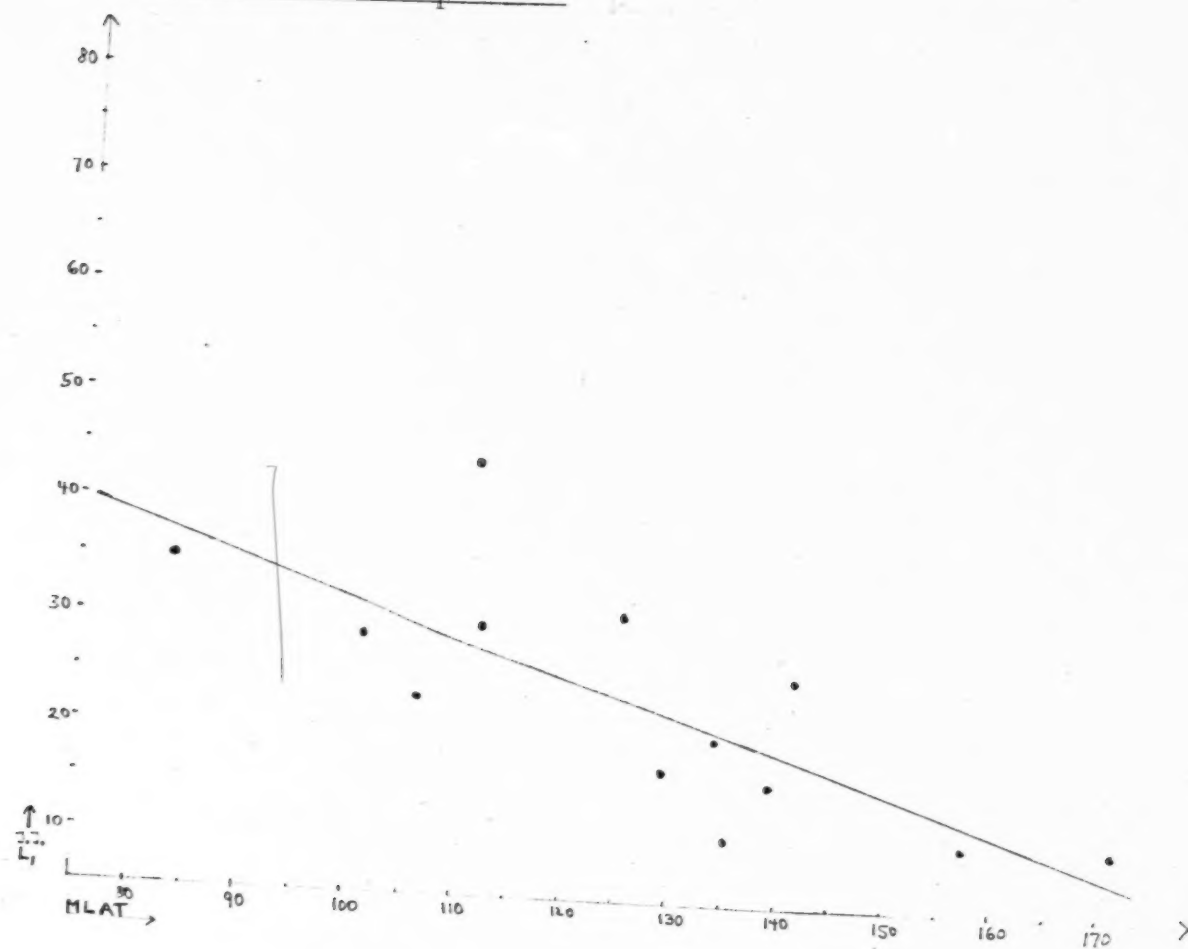
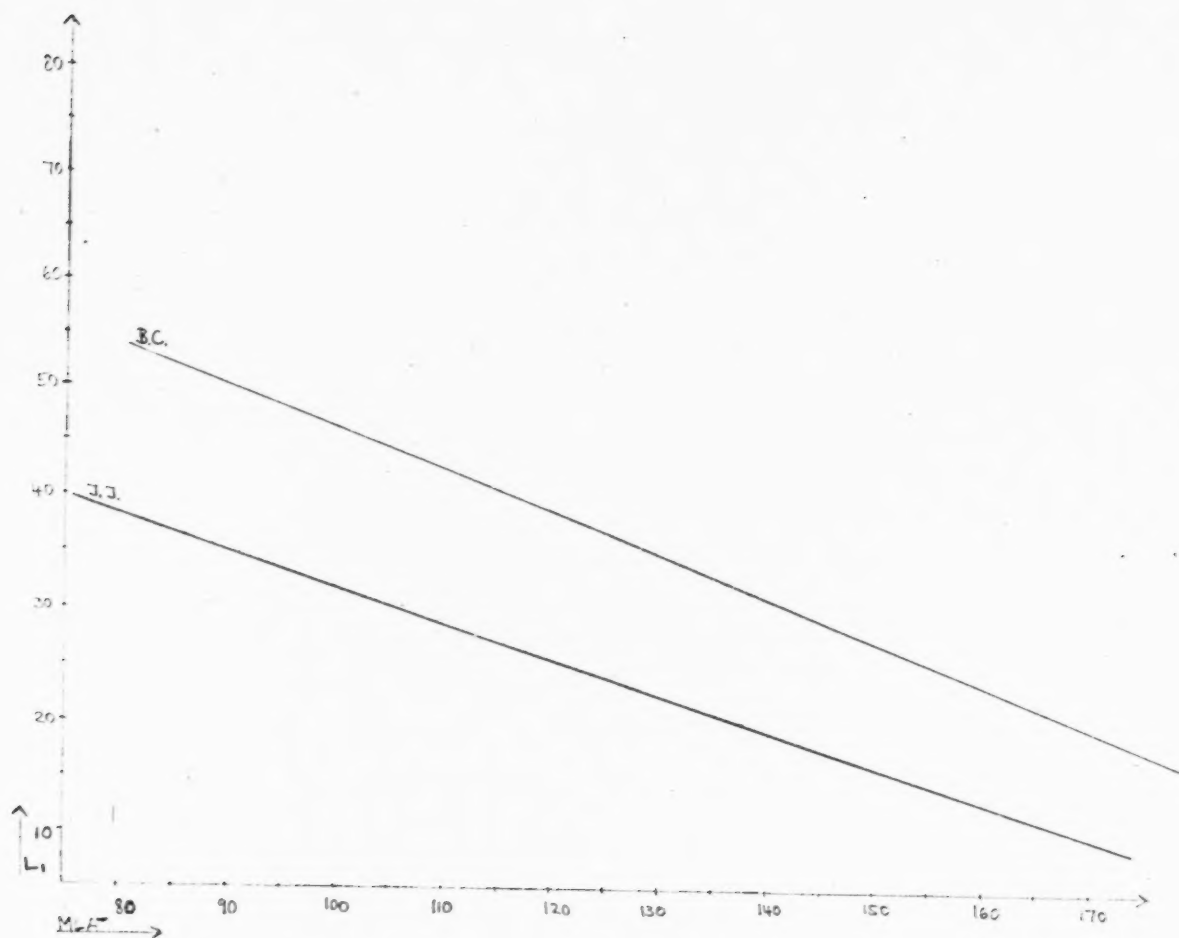
LEVEL 1: Long Tests (L_1) - I.I.

Figure 10

LEVEL 1: Comparison I.I. vs. B.C., L₁

Summary (Figures 8-10)

Whether taking the MLAT-scores into account or not, it can be said that the I.I. students are significantly better than the B.C. students (see a. and e.).

In our analysis it was found that the MLAT is a poor predictor. In comparing t-test scores, with or without using the pretest score, they had about the same values.

There seems to be a greater variation among the B.C. test scores (L_1) than among the I.I. scores. This might be due to better preparation on the part of the I.I. students (c.).

It can also be said that the rate at which the post-test scores change in relation to the MLAT-scores is the same for both I.I. and B.C. students, as can be seen from the equal slopes in the regression line (d.).

Comparison 2:

The two groups, I.I. and B.C., were compared on the basis of their average test score $\bar{U}_I = \frac{U_2 + U_3 + U_4}{3}$. There were 36 observations of $\bar{U}_I$ from B.C. and 20 observations from I.I.

a. The ordinary t-test (without regard to pre-test scores) yields a value of 4.87, which is highly significant.

b. Taking into consideration the pre-test score MLAT, we tried to fit a regression line to the observed values of (MLAT, $\bar{U}_I$) in the I.I. group and the same in group B.C. This was done to explain some of the variation in the $\bar{U}_I$ scores by the variations in the MLAT-scores of the involved students.

For each student a dot was plotted representing his MLAT and $\bar{U}_I$ score. The regression line was then computed and plotted. This was done separately for the

B.C. and I.I. groups.

If the equation of the regression line is

$$D_I = \alpha + \beta (MLAT - \overline{MLAT})$$

β represents the slope of the line, α the intercept of the line at $MLAT = \overline{MLAT}$ ($\overline{MLAT}$ is the average of MLAT-scores for the respective group under consideration).

The following values for α and β were computed, as was s^2 (the residual sum of squares divided by the degrees of freedom), which measures the variation of the data around the fitted line.

	I.I.		B.C.
α	= 2.02	α	= 5.18
β	= -.0361	β	= -.0612
s^2	= .9999	s^2	= 6.5

c. The F-test for comparing the two variations around their respective regression lines shows a significant difference (at level .005), the F-value being 3.25 (degrees of freedom: 34 and 18 respectively).

d. The t-test for equality of slopes gives a t-value of .8217, which is not significant. There is no compelling reason to believe that the two groups are different as far as the slopes of their regression lines are concerned.

e. Considering the slopes equal for I.I. and B.C., one can test the difference of the α value, and here the t-value is 5.07, which is highly significant.

Figure 11

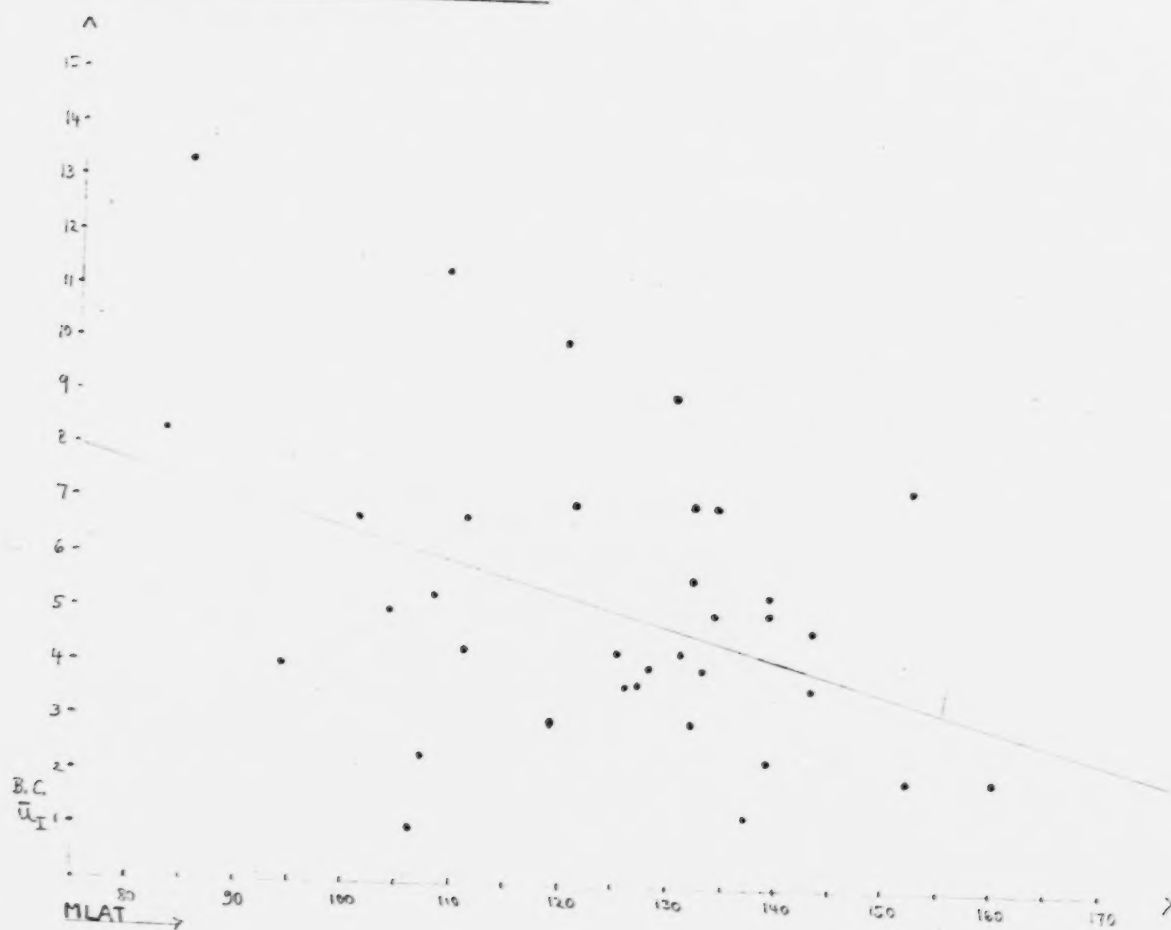
LEVEL 1: Short Tests ($\bar{U}_I$) - B.C.

Figure 12

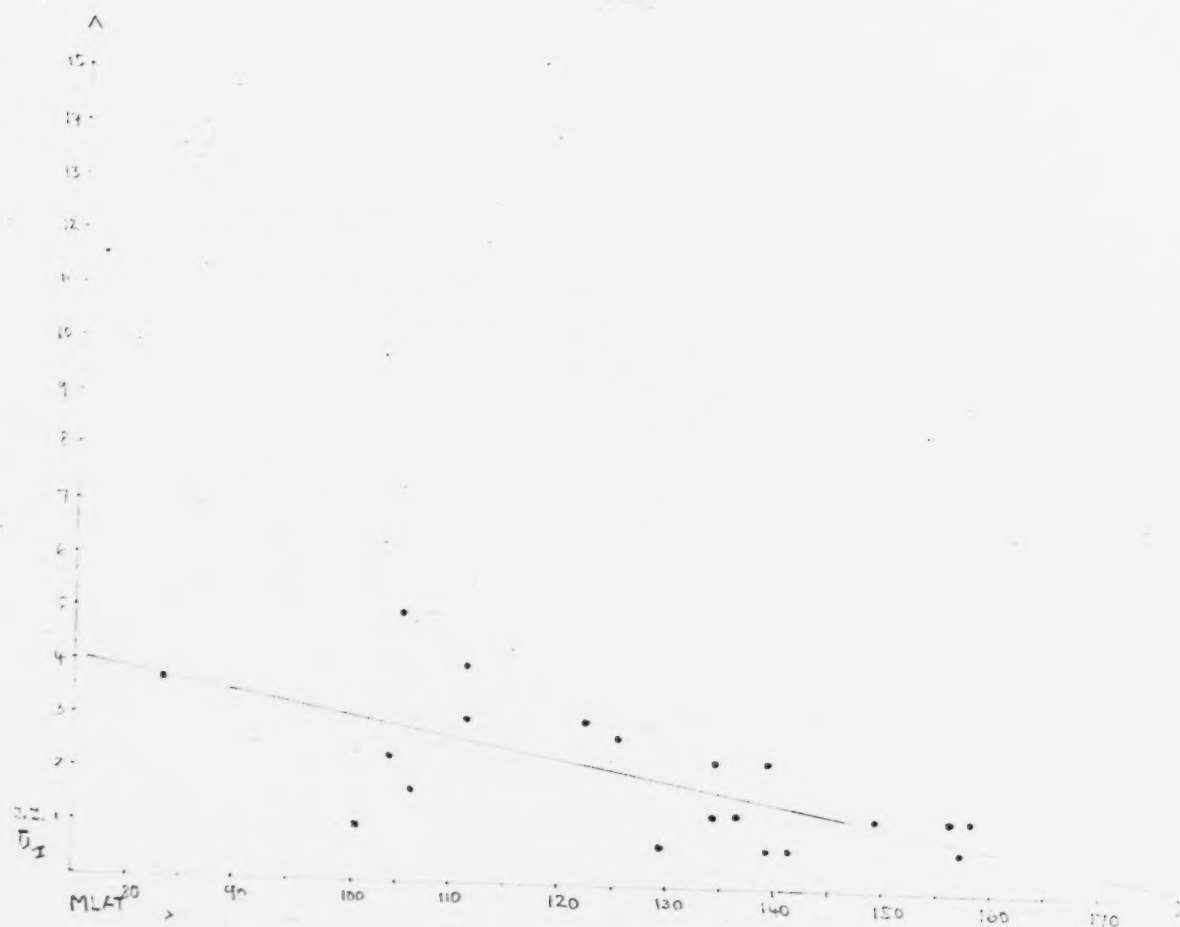
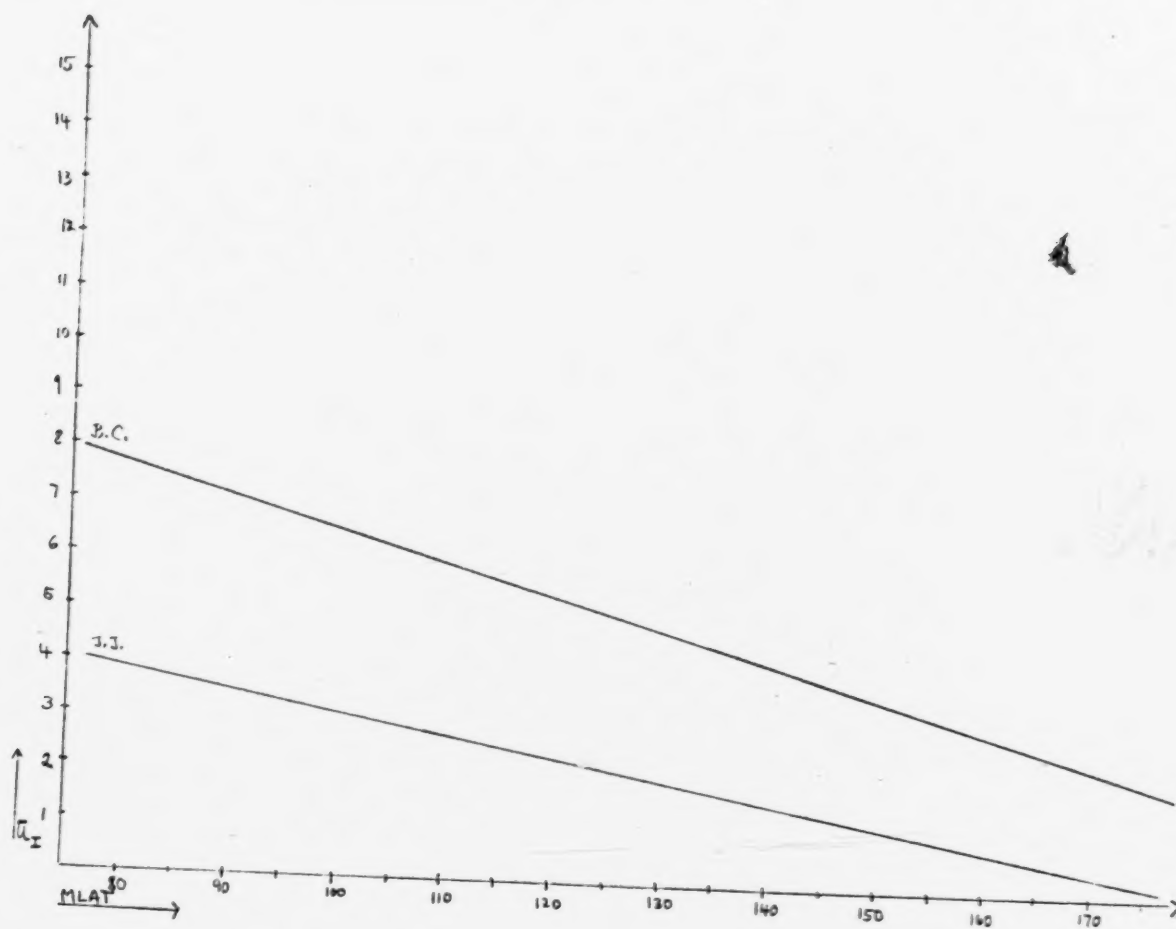
LEVEL 1: Short Tests ($\bar{U}_T$) - I.I.

Figure 13

LEVEL 1: Comparison I.I. vs. B.C. on U_I



Summary (Figures 11-13)

The significant difference between I.I. and B.C. seems to be even more pronounced in the $\bar{U}_I$ scores than in the previous comparison (comparison 1) on L_1 scores (see a. and e.).

There seems to be a much greater variation among the B.C. test scores ($\bar{U}_I$) than among the I.I. scores. This might be due to better preparation on the part of the I.I. students (c.).

Also, it can be said that the MLAT-score affects I.I. and B.C. students at the same rate, equal slopes in regression line (d.).

Comparison 3:

The two groups I.I. and B.C. were compared on the basis of their test scores in L_2 (three-hour test in the second quarter of German). There were 38 observations from B.C. and 15 from I.I.

a. The ordinary t-test (without regard to pretest scores) yields a value of 2.72, which is significant at level .005 (one-sided).

b. Taking into consideration the pretest score MLAT, we tried to fit a regression line to the observed values of (MLAT, L_2) in the I.I. group and the same in group B.C. This was done to explain some of the variation in the L_2 scores by the variations in the MLAT scores of the involved students.

For each student a dot was plotted representing his MLAT and L_2 scores. The regression line was then computed and plotted. This was done separately for the B.C. and I.I. group. If we assume the equation of the regression line to be:

$$L_2 = \bar{X} + Y (MLAT - \overline{MLAT})$$

δ represents the slope of the line, X the intercept of the line at $\overline{\text{MLAT}}$ ($\overline{\text{MLAT}}$ is the average of MLAT-scores for the respective group under consideration).

The following values for X and δ were computed, as was s^2 (the residual sum of squares divided by the degrees of freedom), which measures the variation of the data around the fitted line.

<u>I.I.</u>	<u>B.C.</u>
$X = 27.46$	$X = 42.08$
$\delta = -.4130$	$\delta = -.1573$
$s^2 = 203.24$	$s^2 = 335.35$

c. Here the F-test for comparing the two variations around their respective regression lines shows no significant difference, the F-value being 1.65 (degrees of freedom: 36 and 11).

d. The t-test for equality of slopes gives a t-value of .80 (49 degrees of freedom), which is not significant.

There is no compelling reason to believe that the two groups are different as far as slopes are concerned.

e. Considering the slopes equal for I.I. and B.C., one can test the difference of the X value, and here the t-value is 2.19 (50 degrees of freedom), which is significant at level .05.

Figure 14

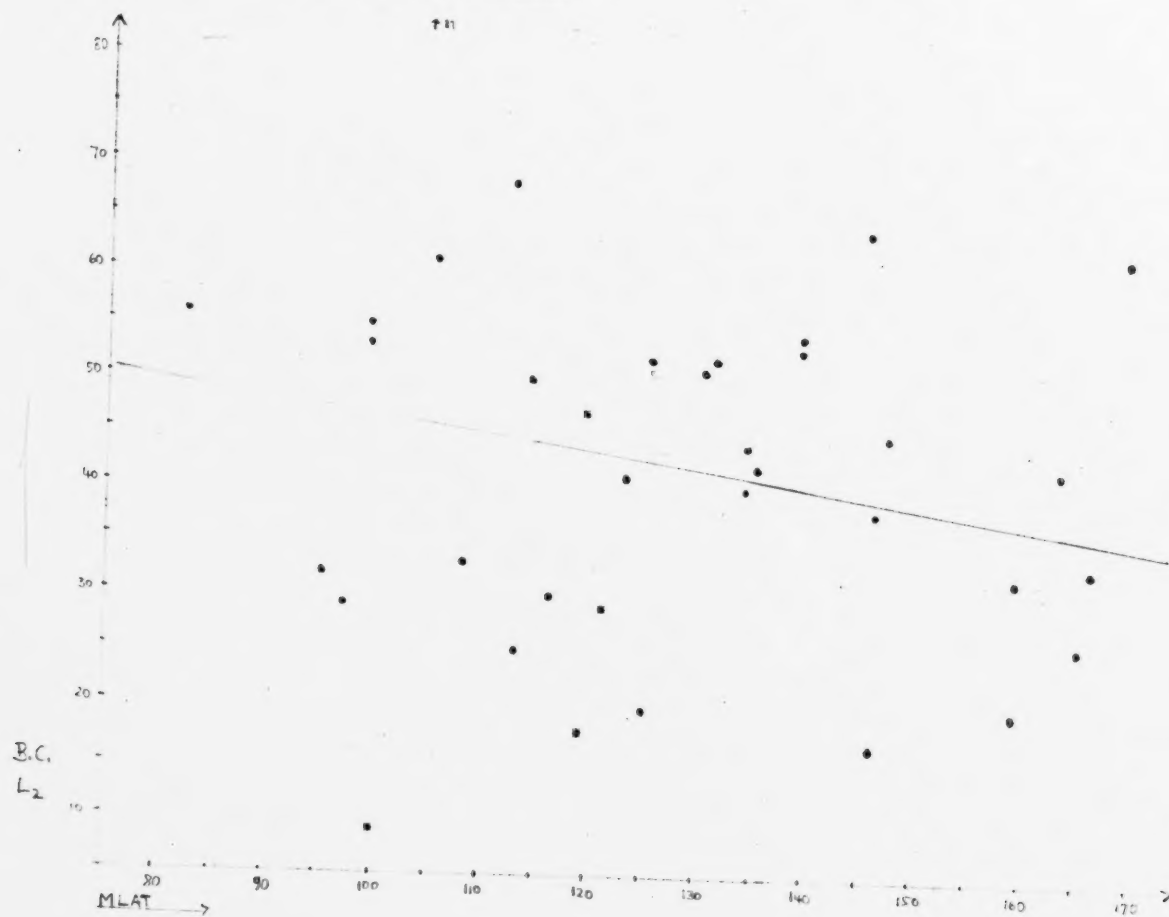
LEVEL 2: Long Tests (L_2) - B.C.

Figure 15

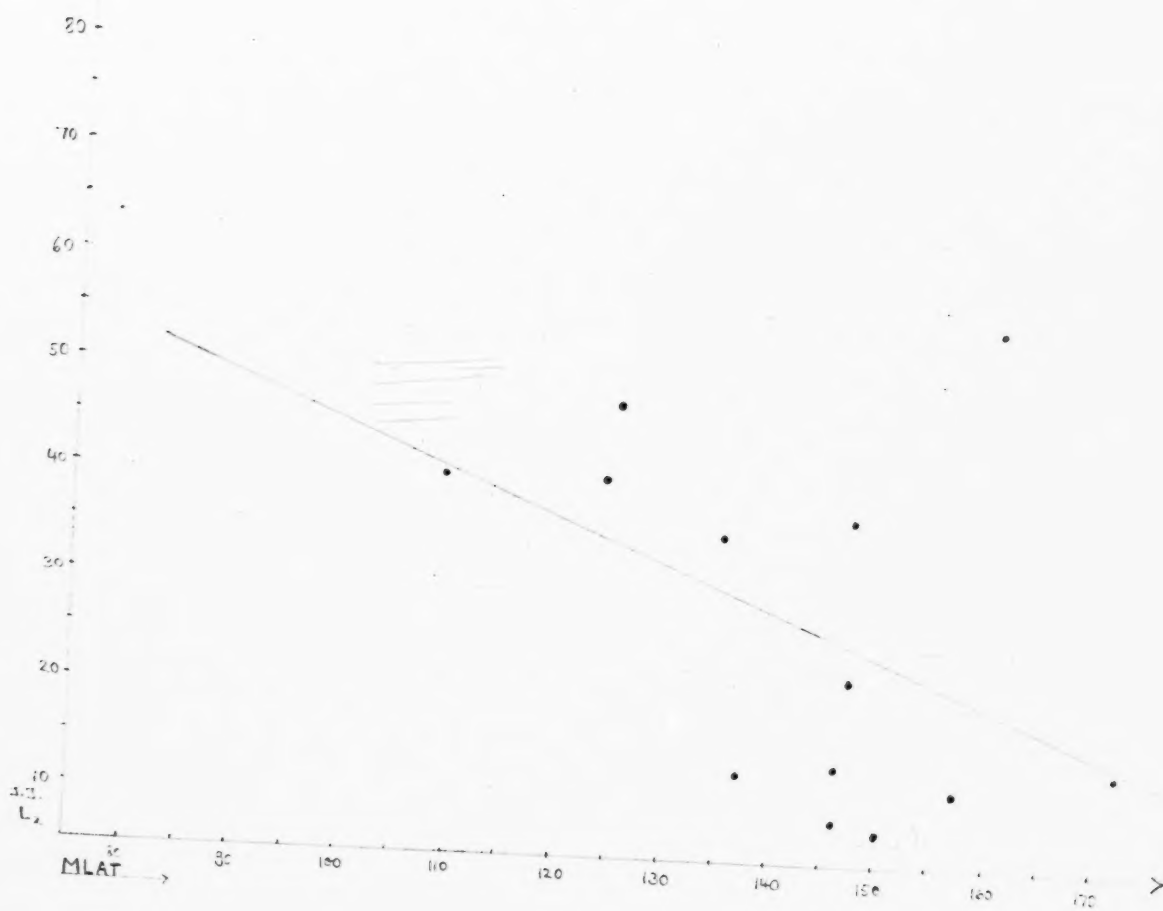
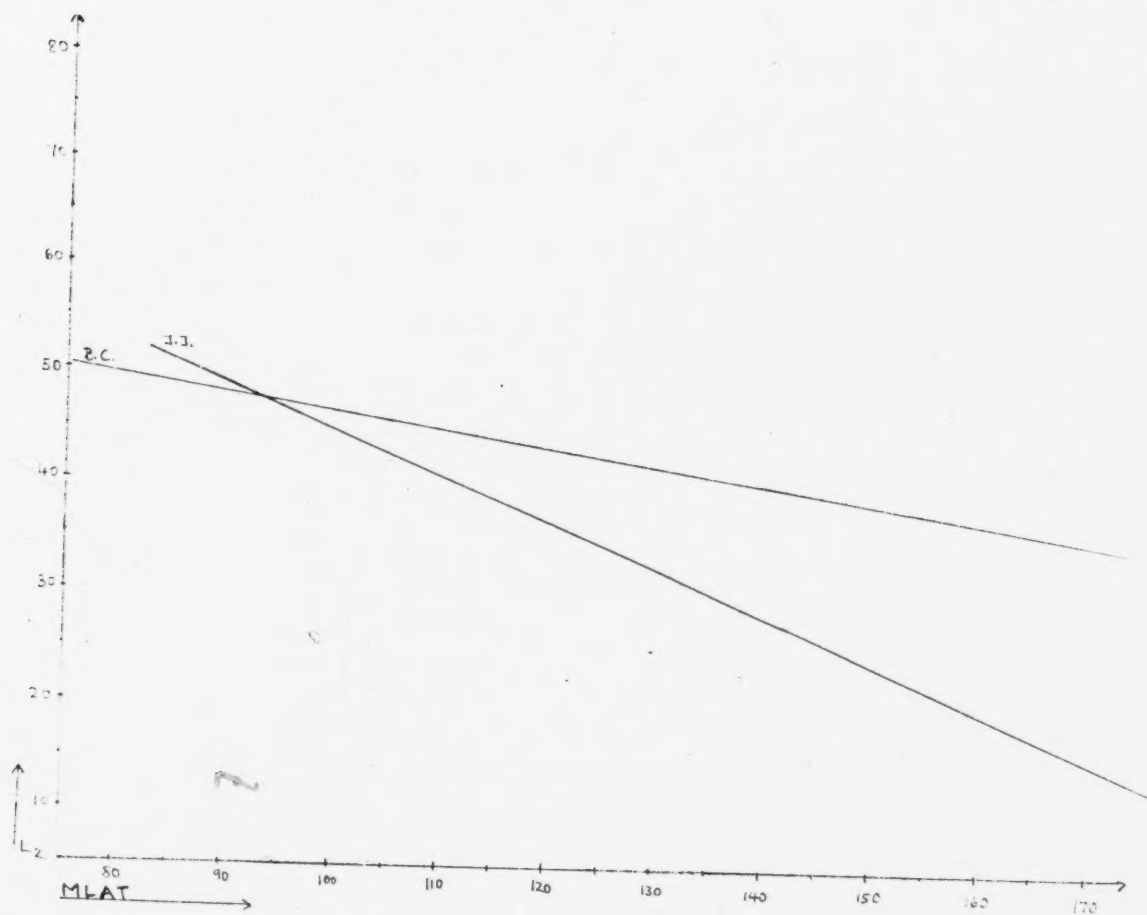
LEVEL 2: Long Tests (L_2) - I.I.

Figure 16

LEVEL 2: Comparison I.I. vs. B.C. on L_2



Summary (Figures 14-16)

Whether the MLAT-scores are taken into account or not, it can be said that the I.I. students are significantly better than the B.C. students (see a. and e.).

It should be noted that the t-value obtained in e. by taking into consideration the MLAT-scores is not as significant as the t-value in a. This may be due to the fact that the I.I. group on this level had subjects with generally higher MLAT-scores, so that some of that difference between B.C. and I.I. can be attributed to the difference in MLAT-scores in these two groups. The variation in the B.C. and I.I. scores (L_2) seem to be the same (c.).

It can also be said that the MLAT-scores affect the I.I. and B.C. students at about the same rate, although in the graph there seems to be some optical difference of slope in the two lines, which, however, is not statistically significant (d.). From the graph it may be seen that MLAT has little prediction value for L_2 , especially for the B.C. students.

Comparison 4:

The two groups I.I. and B.C. were compared on the basis of $\bar{U}_{II} = \frac{U_6 + U_7 + U_8}{3}$

their average test score in three thirty-minute tests at the end of each completed unit in Level 2, or during the second quarter of German. There were 30 observations from B.C. and 10 from I.I.

a. The ordinary t-test (without regard to pretest scores) yields a value of 5.10, which is highly significant, i.e., if the two groups are equal, then the chance of observing such a large t-value is very rare.

b. Taking into consideration the pre-test score MLAT, we tried to fit a regression line to the observed values of (MLAT, $\bar{U}_{II}$) in the I.I. group and the

same in group B.C. This was done to explain some of the variation in the $\bar{U}_{II}$ scores by the variations in the MLAT-scores of the involved students. For each student a dot was plotted representing his MLAT and $\bar{U}_{II}$ scores. Then the regression line was computed and plotted. This was done separately for the B.C. and I.I. groups. If we assume the equation of the regression line to be:

$$\bar{U}_{II} = \lambda + \gamma (\text{MLAT} - \overline{\text{MLAT}})$$

γ represents the slope of the line, λ the intercept of the line at $\text{MLAT} = \overline{\text{MLAT}}$ ($\overline{\text{MLAT}}$ is the average of MLAT-scores for the respective group under consideration).

The following values for λ and γ were computed, as was s^2 (the residual sum of squares divided by the degrees of freedom), which measures the variation of the data around the fitted line.

<u>I.I.</u>	<u>B.C.</u>
$\lambda = 3.5$	$\lambda = 9.22$
$\gamma = -.0226$	$\gamma = .0086$
$s^2 = 2.318$	$s^2 = 12.071$

c. The F-test comparing the two variations around their respective regression lines yields an F-value of 5.2 (degrees of freedom: 28 and 8), which could be considered significant at level .01.

d. The t-test for equality of slopes gives a t-value of .169, which is not significant.

e. Considering the slopes equal for I.I. and B.C., one can test the difference of the λ value, and here the t-value is 4.75 (37 degrees of freedom), which is highly significant.

Figure 17

LEVEL 2: Short Tests ($\bar{U}_{II}$) - B.C.

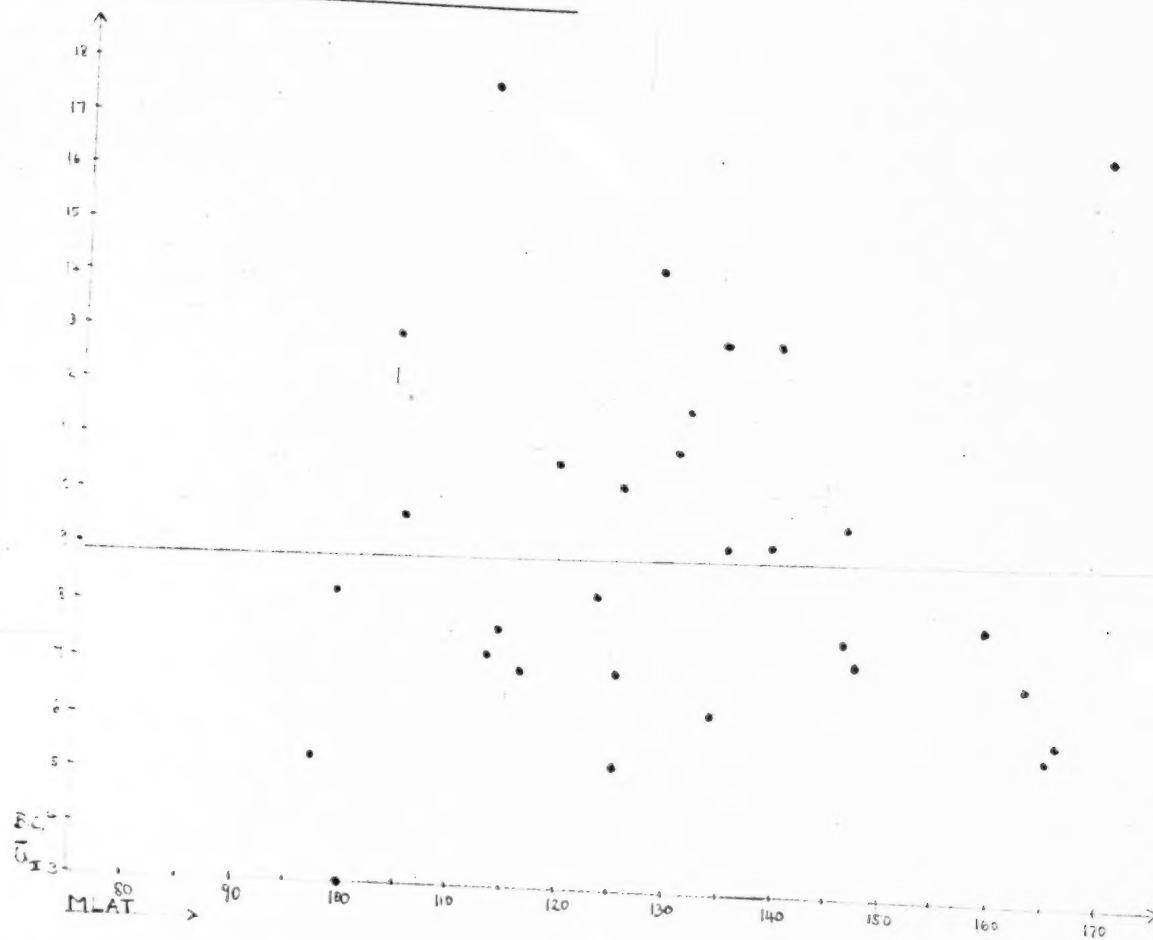


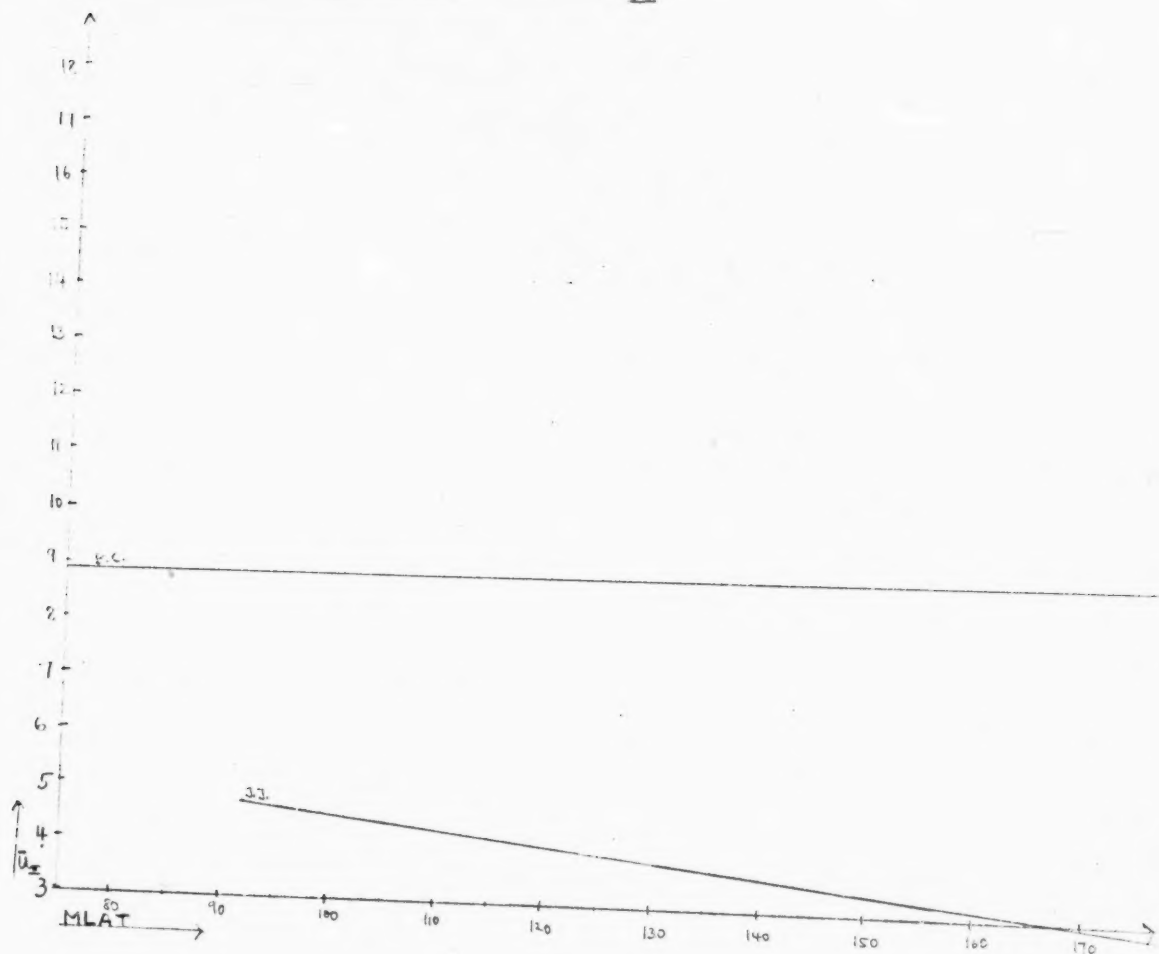
Figure 18

LEVEL 2: Short Tests (U_{II}) - I.I.



Figure 13

LEVEL 2: Comparison I.I. vs. B.C. on U_{II}



Summary (Figures 17-19)

The significant difference between I.I. and B.C. seems to be more pronounced than in Comparison 3 (Figure 16). Comparisons 2 (Figure 10) and 4 (Figure 13) present similar evidence showing that the difference in performance between the I.I. and B.C. groups is more pronounced on the short tests. This may, in part, be due to the fact that the I.I. student determines when he takes a unit examination, whereas the B.C. examination times are scheduled by the instructor. Furthermore the I.I. students prepare themselves by completing Practice Tests.

The variation seems to be larger among the B.C. test scores (c.). Although there seems to be an optical difference in slopes, this is not significant. In the graph for the B.C. students, however, it seems to suggest that the MLAT-score is a poor predictor for the $\bar{U}_{II}$ -scores.

Comparison 5:

It did not seem statistically worthwhile to investigate the third-quarter German students in the same way as above. The small sample of students (28) did not warrant a detailed investigation. We therefore computed the average of mistakes for each posttest L_3 , U_{11} , U_{13} and U_{14} . This was done separately for the I.I. and B.C. groups, and we compared them. Not considering the MLAT-scores for both I.I. and B.C., we can, however, conclude the following:

The average number of minus points of the B.C. students on L_3 (the three-hour test given at the end of the third quarter of German) is 44.87, whereas the average for the I.I. is 31.20. This cannot be considered significant in view of the small number of observations obtained (B.C. 23 observations, I.I. 5 observations) and because of the variation within the groups.

The analysis (not considering MLAT-scores) again verifies the phenomenon observed with the short tests U_I and U_{II} , namely that I.I. students make fewer mistakes and are better prepared. The average minus points are given below:

U_{11} : B.C. -8.05 I.I. -4.72

U_{13} : B.C. -10.05 I.I. -7.66

U_{14} : B.C. -8.32 I.I. -3.27

C. Evaluation of Oral Performance

Since our initial appraisal of the oral performance of students in both tracks yielded inconclusive results, a detailed examination of oral work was conducted for this second appraisal. The same oral questions of various types were asked of both B.C. and I.I. students at appropriate times during the quarter, and tape recordings were made of the answers. From the total sample of several hundred taped oral performances, six questions were randomly selected (two for each level), and then B.C. and I.I. performances were randomly intersplined for each of the three levels, yielding a tape of 40 performances, 15 each for Levels 1 and 2 and 10 for Level 3. Fourteen judges (faculty members and teaching assistants who were not instructors of the students tested) were then asked to listen to the performances of all 40 students and to evaluate them, giving each student a score along a scale from 10 to 0 (10 = outstanding, 5 = average, 0 = extremely poor) on each of four criteria: pronunciation, vocabulary, fluency, grammar. The results, summarized below, show that with one exception there was no significant difference between the I.I. and B.C. groups on any level on any of the criteria. We therefore conclude that the oral performance of I.I. students is as good as that of B.C. students.

Table 1: Oral Performance, Sample t-tests, I.I. vs. B.C.

	N	Pronunciation	Vocabulary	Fluency	Grammar
Level I	I.I. = 8 B.C. = 7	t = .41	t = .63	t = 1.6	t = .09
Level II	I.I. = 8 B.C. = 7	t = .0047	t = .643	t = .419	t = .58
Level III	I.I. = 4 B.C. = 6	t = .51	t = .21	t = .14	t = .16

These t-values show no significant difference in each case, except that $t = 1.6$ can be considered significant at the .1 level.

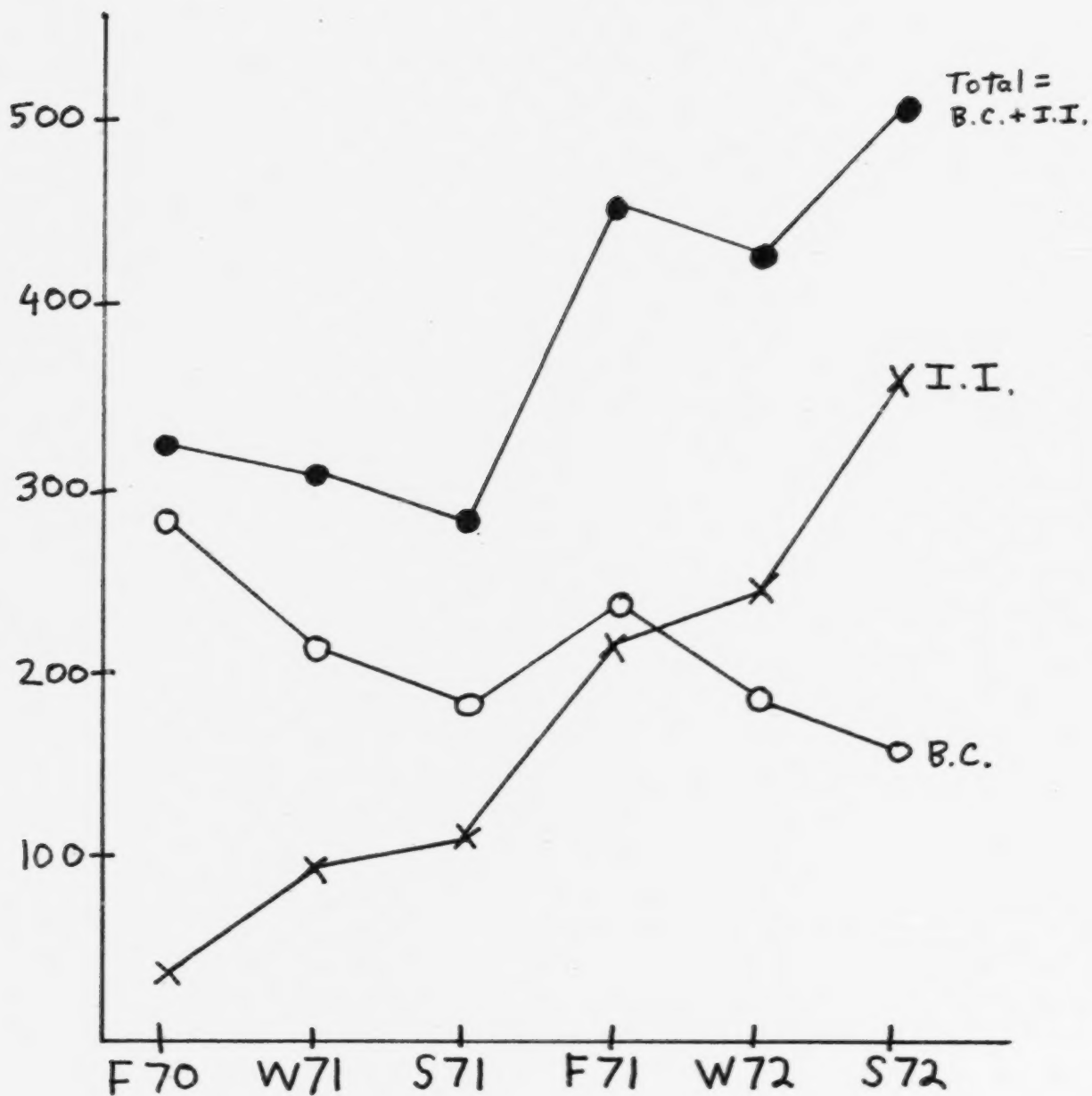
III. Other Measures and Results

A. Enrollment. When the course was first begun in the Fall Quarter of 1970, 37 students elected to be in the I.I. program. At that time 290 students were enrolled in the B.C. Since that time enrollments in the I.I. course have steadily increased, whereas enrollment in the B.C. has somewhat decreased (see Figure 20). Fluctuations in B.C. enrollment from Fall to Spring of each year conform to the normal pattern at Berkeley.

It may be noted that the total combined enrollment of B.C. and I.I. increased substantially due to the increased I.I. enrollment.

FIGURE 20:

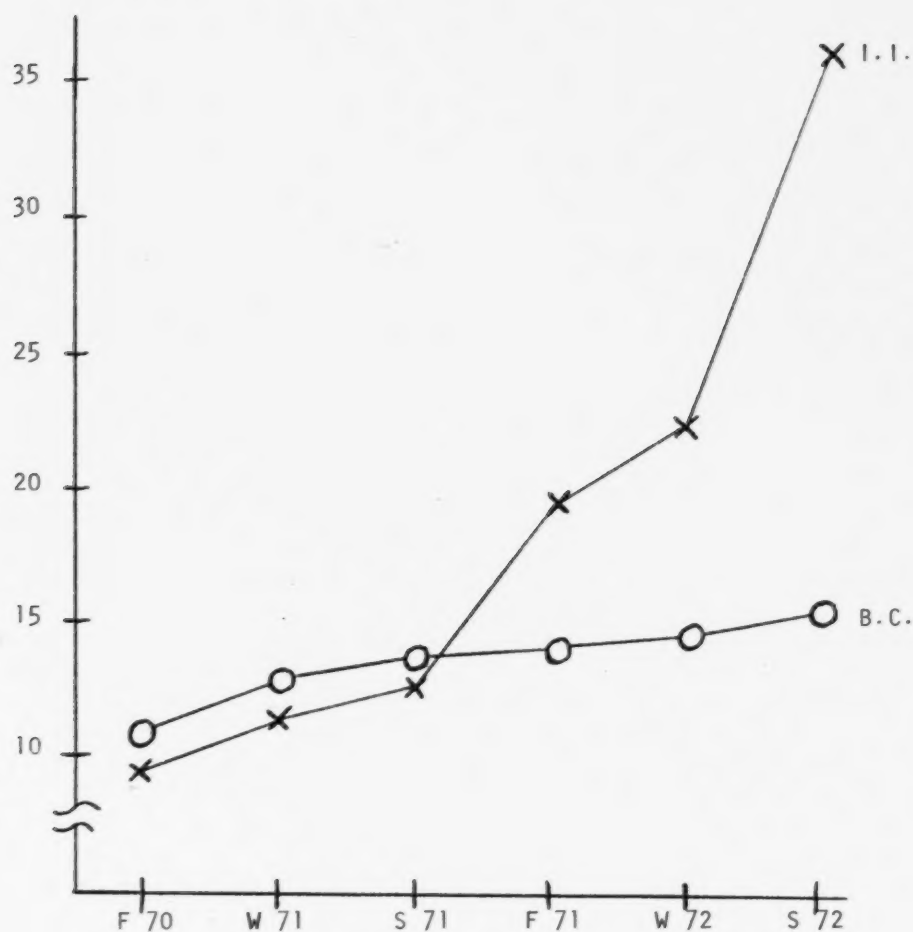
Enrollment, I.I./B.C.



B. Student-Instructor Ratios. It should be noted (Figure 21) that I.I. section size has increased from 9.8 to 35.7 students per instructor, that is, it has more than tripled, while B.C. section size increased by about 50% over the last two years.

Figure 21: Average No. of Students per Instructor

	I.I.	B.C.
F 70	9.8	10.6
W 71	11.5	12.6
S 71	12.2	12.8
F 71	19.5	13.4
W 72	22.2	14.2
S 72	35.7	16.2

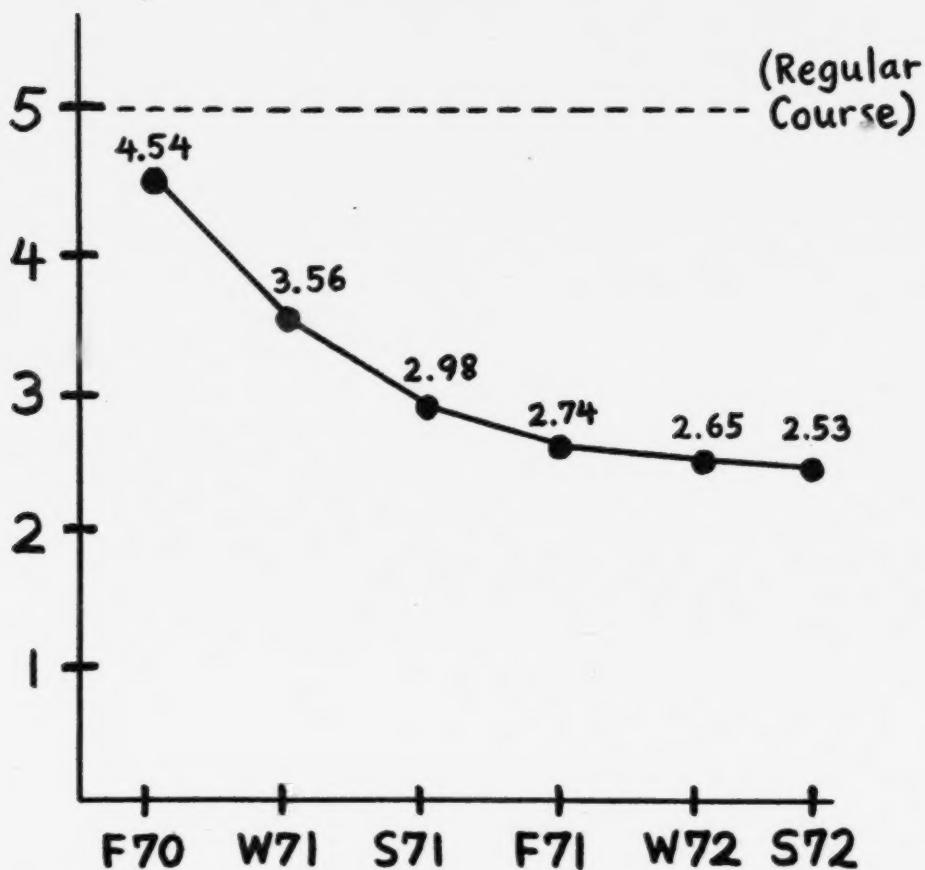


The increase in student-instructor ratio for the B.C. is primarily due to the stringency of financial resources in the University. The additional increase in student-instructor ratio in the I.I. course is, however, primarily due to the fact that students, as was mentioned earlier, need to contact an instructor approximately only one-third as often as is necessary in the B.C.

C. Achievement of Units. It may be noted from Figure 22 that the average number of units achieved by students in the I.I. program has decreased from 4.54 in the Fall of 1970 to 2.53 in the Spring of 1972.

FIGURE 22:

Average No. of Quarter Units Achieved per Student



The units achieved per quarter now seems to be leveling off and to amount to approximately half the units achieved in the B.C. It was noted in B. above (Figure 21) that the ratio of students per instructor in the I.I. program increased rather dramatically. Not only is this due to the fact that students need to have contacts with instructors less often because of the instructional mode, but we now see from the results displayed in Figure 22 that it is also due to the fact that the average number of units being taken by I.I. students is only approximately half of the number of units taken by students in the regular course. Although the number of contacts an I.I. student has with the instructor, as mentioned earlier, is approximately one-third, the total time each instructor is scheduled to teach is approximately the same, since in this program many contacts are on a one-to-one basis. In our opinion this more personal contact of individual students with the instructor and the fact that students and instructors address themselves to the specific problems of each individual student account in part for the fact that achievement on the tests is generally higher for the I.I. students.

D. Workload. An analysis of students' study lists, including courses other than I.I., has shown that students are able to balance their programs in such a way that their quarterly workload in I.I. is well integrated with their other courses. As a general rule it was observed that students who take a heavy load of I.I. units also have a higher total number of units on their quarterly program.

E. Student and Instructor Questionnaires. As during the first year, we prepared student and instructor questionnaires. The questionnaires of the second year were similar to those of the first year. An analysis of the answers

shows no significant differences from the responses obtained in the first year's study. Readers of this report who are interested in the results of the questionnaire are asked to contact the authors. We feel that the inclusion of the questionnaire results would unduly lengthen this paper, and we have therefore decided not to include them.

IV. Conclusion

The evaluation of student achievement in the second year shows significantly better performance by Individualized Instruction students in reading and writing and equal oral performance compared with the Basic Course students. These findings confirm the conclusions of our first study and substantiate them with more rigorous statistical evidence. The data gathered over two years from a substantial student population seems to demonstrate the basic soundness of individualized foreign language instruction. We recommend its use in a variety of instructional settings as well as its further refinement.